UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/662,985	09/15/2003	Min Chu	M61.12-0565	2246

27366 7590 04/06/2009

WESTMAN CHAMPLIN (MICROSOFT CORPORATION)
SUITE 1400
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402

EXAMINER	
COLUCCI, MICHAEL C	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
04/06/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/662,985	CHU ET AL.	
	Examiner	Art Unit	
	MICHAEL C. COLUCCI	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *02 February 2009*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *23,25,26,28,29 and 31-36* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *23,25,26,28,29 and 31-36* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 02/02/2009 has been entered.

Response to Arguments

2. Applicants arguments with respect to claims 23, 25, 26, 28, 29, and 31-36 have been considered but are moot in view of the new grounds of rejection. After consideration of the Remarks filed 02/02/2009 as well as the claims in light of the specification, Examiner has withdrawn "Recent improvements on Microsoft's trainable text-to-speech system-Whistler" (hereinafter Huang) and instead incorporated Coorman et al. US 6665641 B1 (hereinafter Coorman) for the rejection of claim 33 and similarly claims 23 and 28. Though the teachings of Huang appear to inherently teach concatenation with respect to *higher level* order, wherein higher level order merely refers to more than one criteria of speech identification (i.e. phoneme position, phoneme stress, and phoneme pitch, duration, etc.). Examiner has incorporated Coorman to explicitly address these issues, wherein Coorman like the present invention teaches various speech portions and multiple identification methods that are weighted based on

a distance measure that are dependent on various factors for a speech concatenation candidate (i.e. stress, phoneme position, etc.). Further, Coorman teaches an improvement of various well known speech segment identification techniques.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 23, 25, 26, 28, 29, 31, and 32 are rejected under 35 U.S.C. 103(a) as being unpatentable over Coorman et al. US 6665641 B1 (hereinafter Coorman) in view of Seide US 5857169 A (hereinafter Seide).

Re claims 23, 28, and 33 , Coorman teaches a method of selecting speech segments for concatenative speech synthesis the method comprising:

parsing an input text into speech units (Col. 2 line 58 – Col. 3 line 12 & Fig. 1)

identifying context information for each speech unit based on its location in the input text and at least one neighboring speech unit (Col. 3 lines 19-54).

identifying a set of candidate speech segments for each speech unit based on the context information, wherein identifying a set of candidate speech segments for a speech unit comprises applying the context information for a speech unit (Col. 3 lines

19-54) to a decision tree to identify a leaf node containing candidate speech segments for the speech unit,

wherein identifying the sequence of speech segments comprises using an objective measure comprising a plurality of components, each component having an associated weighing value (Col. 16 lines 47-67), and wherein a first component is based on one factor in the set of factors below, and a second component is a combination of at least two factors from the set of factors, the set of factors including

an indication of a position of a speech unit in a phrase (Col. 3 lines 19-54, utterance);

an indication of a position of a speech unit in a word (Col. 3 lines 19-54, word);

an indication of a category for a phoneme preceding a speech unit (Col. 3 lines 19-54, left and right context);

an indication of a category for a phoneme following a speech unit (Col. 3 lines 19-54, left and right context);

an indication of a category for tonal identity of the current speech unit;

an indication of a category for tonal identity of a preceding speech unit;

an indication of a category for tonal identity of a following speech unit;

an indication of a level of stress of a speech unit (Col. 3 lines 19-54, stress markers);

an indication of a coupling degree of pitch, duration and/or energy with a neighboring unit (Col. 3 lines 19-54, exact duration);

an indication of a degree of spectral mismatch with a neighboring speech unit.

identifying a sequence of speech segments from the candidate speech segments; and

generating synthesized speech using the sequence of speech segments without further prosody modification

However Coorman fails to particularly teach to identify a leaf node containing candidate speech segments for the speech unit

Seide teaches a localizer 50 performs the locating by, for each observation vector, searching the tree structure corresponding to a reference unit until at the lowest tree level a number of leaf nodes are selected. For the selected leaf nodes, the localizer 50 determines how well the observation vector matches this reference unit. This involves for each selected leaf node using the reference probability density, which corresponds to the leaf node, to calculate an observation likelihood for the observation vector. For each reference unit, the observation likelihoods, which have been calculated for one observation vector, are combined to give a reference unit similarity score. For each reference pattern, the reference unit similarity scores of the reference unit, which correspond to the reference pattern are combined to form a pattern similarity score. This is repeated for successive observation vectors. The reference pattern for which an optimum, such as a maximum likelihood, is calculated for the pattern similarity score is located as the recognized pattern. The description focuses on locating reference probability densities and calculating observation likelihoods. It is well understood in the art how this key element can be used in combination with other

techniques, such as Hidden Markov Models, to recognize a time sequential pattern, which is derived from a continual physical quantity. It is also well understood in the art how techniques, such as a leveled approach, can be used to recognize patterns which comprise a larger sequence of observation vectors than the reference patterns. For instance, it is known how to use sub-word units as reference patterns to recognize entire words or sentences. It is also well understood how additional constraints, such as a pronunciation lexicon and grammar, may be placed on the pattern recognition. The additional information, such as the pronunciation lexicon, can be stored using the same memory as used for storing the reference pattern database (Seide Col. 8 lines 31-67).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Coorman to incorporate identifying a leaf node containing candidate speech segments for the speech unit as taught by Seide to allow for an optimized output of natural sounding speech based on prosodic, lexical, and syntactical features as well as grammatical analysis to produce the highest matching score (Seide Col. 8 lines 31-67).

Re claims 25, 29, 31, 34, and 35, Coorman teaches the method of claim 23 wherein identifying a set of candidate speech segments further comprises pruning some speech segments (Col. 4 lines 18-30) from a leaf node based on differences between the context information of the speech unit from the input text and context information associated with the speech segment

However, Coorman fails to teach pruning some speech segments from a leaf node

Seide teaches a localizer 50 performs the locating by, for each observation vector, searching the tree structure corresponding to a reference unit until at the lowest tree level a number of leaf nodes are selected. For the selected leaf nodes, the localizer 50 determines how well the observation vector matches this reference unit. This involves for each selected leaf node using the reference probability density, which corresponds to the leaf node, to calculate an observation likelihood for the observation vector. For each reference unit, the observation likelihoods, which have been calculated for one observation vector, are combined to give a reference unit similarity score. For each reference pattern, the reference unit similarity scores of the reference unit, which correspond to the reference pattern are combined to form a pattern similarity score. This is repeated for successive observation vectors. The reference pattern for which an optimum, such as a maximum likelihood, is calculated for the pattern similarity score is located as the recognized pattern. The description focuses on locating reference probability densities and calculating observation likelihoods. It is well understood in the art how this key element can be used in combination with other techniques, such as Hidden Markov Models, to recognize a time sequential pattern, which is derived from a continual physical quantity. It is also well understood in the art how techniques, such as a leveled approach, can be used to recognize patterns which comprise a larger sequence of observation vectors than the reference patterns. For instance, it is known how to use sub-word units as reference patterns to recognize

entire words or sentences. It is also well understood how additional constraints, such as a pronunciation lexicon and grammar, may be placed on the pattern recognition. The additional information, such as the pronunciation lexicon, can be stored using the same memory as used for storing the reference pattern database (Seide Col. 8 lines 31-67).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Coorman to incorporate pruning some speech segments from a leaf node as taught by Seide to allow for an optimized output of natural sounding speech based on prosodic, lexical, and syntactical features as well as grammatical analysis to produce the highest matching score (Seide Col. 8 lines 31-67).

Re claims 26, 32, and 36, Coorman teaches the method of claim 23 wherein identifying a sequence of speech segments comprises using a smoothness cost (Col. 11 lines 35-54, transition cost that scores 'joinability') that is based on whether two neighboring candidate speech segments appeared next to each other in a training corpus (Col. 2 lines 38-49, system that learns).

Conclusion

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. US 4979216 A, US 6366883 B1, US 5715367 A.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626

(571)-270-1847
Michael.Colucci@uspto.gov

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626